

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Gentry, Daniel
Lonsdale, John
Payne, David
Pearson, Stewart
Van Aller, Glenn

(ii) TITLE OF THE INVENTION: Novel FabD

(iii) NUMBER OF SEQUENCES: 6

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: SmithKline Beecham Corporation
(B) STREET: 709 Swedeland Road
(C) CITY: King of Prussia
(D) STATE: PA
(E) COUNTRY: USA
(F) ZIP: 19406-0939

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE: 14-NOV-1997
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 60/031,160
(B) FILING DATE: 18-NOV-1996

09300397-054800

Sub
A1

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Gimmi, Edward R
(B) REGISTRATION NUMBER: 38,891
(C) REFERENCE/DOCKET NUMBER: P50593

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 610-270-4478
(B) TELEFAX: 610-270-5090
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 921 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGACTAAAA CAGCCTTTT ATTTGCTGGT CAAGGTGCCC AGTATCTAGG GATGGGACGG 60
GATTTCTATG ATCAGTATCC GATTGTCAAA GAAACGATTG ATCGAGCGAG TCAGGTGCTC 120
GGTTATGATT TACGTTATCT CATCGATACG GAAGAAGACA AACTCAATCA GACCCGCTAT 180
ACGCAACCAG CCATTCTAGC GACTTCGGTT GCTATCTACC GTTTATTGCA AGAAAAGGGC 240
TATCAGCCTG ATATGGTTGC TGGTTTGTCT CTTGGAGAAT ACTCTGCCTT GGTGGCAAGC 300
GGCGCCTTGG ATTTTGAAGA TCGGTTGCC TTGGTAGCTA AGCGTGGAGC CTATATGGAA 360
GAAGCGGCTC CTGCTGACTC TGGCAAGATG GTAGCAGTTC TCAATACGCC AGTAGAGGTC 420
ATTGAAGAAG CCTGTCAAAA AGCTTCTGAA CTTGGAGTGG TTAATCCAGC CAACTATAAC 480
ACACCTGCAC AAATCGTCAT TGCTGGAGAA GTGGTTGCAG TTGATCGAGC GGTGAACTT 540
TTGCAAGAAG CAGGTGCCAA ACGCTTGATT CCTCTTAAGG TGTCAGGTCC CTTTCACACC 600
GCTCTCCTTG AGCCTGCTAG CCAGAACTA GCTGAACTC TAGCTCAGGT AAGTTTTTCA 660
GATTTTACTT GTCCCCTAGT CGGCAATACA GAAGCTGCTG TGATGCAAAA AGAGGACATT 720
GCTCAGCTCT TGACGCGTCA GGTCAAGGAA CCCGTTTCGT TCTATGAAAG TATTGGGGTC 780
ATGCAAGAAG CAGGCATAAG CAACTTTATC GAGATTGGAC CGGGGAAAGT CTTGTCAGGT 840

TTTGTTAAAA AAATTGATCA AACTGCTCAC TTAGCTCATG TGGAAGATCA AGCGAGTTTA
GTAGCACTTT TAGAAAAATA G

900

921

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 306 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Thr	Lys	Thr	Ala	Phe	Leu	Phe	Ala	Gly	Gln	Gly	Ala	Gln	Tyr	Leu
1				5					10					15	
Gly	Met	Gly	Arg	Asp	Phe	Tyr	Asp	Gln	Tyr	Pro	Ile	Val	Lys	Glu	Thr
			20					25					30		
Ile	Asp	Arg	Ala	Ser	Gln	Val	Leu	Gly	Tyr	Asp	Leu	Arg	Tyr	Leu	Ile
		35					40					45			
Asp	Thr	Glu	Glu	Asp	Lys	Leu	Asn	Gln	Thr	Arg	Tyr	Thr	Gln	Pro	Ala
	50					55					60				
Ile	Leu	Ala	Thr	Ser	Val	Ala	Ile	Tyr	Arg	Leu	Leu	Gln	Glu	Lys	Gly
65					70					75				80	
Tyr	Gln	Pro	Asp	Met	Val	Ala	Gly	Leu	Ser	Leu	Gly	Glu	Tyr	Ser	Ala
			85						90					95	
Leu	Val	Ala	Ser	Gly	Ala	Leu	Asp	Phe	Glu	Asp	Ala	Val	Ala	Leu	Val
		100						105					110		
Ala	Lys	Arg	Gly	Ala	Tyr	Met	Glu	Glu	Ala	Ala	Pro	Ala	Asp	Ser	Gly
		115					120					125			
Lys	Met	Val	Ala	Val	Leu	Asn	Thr	Pro	Val	Glu	Val	Ile	Glu	Glu	Ala
	130					135						140			
Cys	Gln	Lys	Ala	Ser	Glu	Leu	Gly	Val	Val	Thr	Pro	Ala	Asn	Tyr	Asn
145					150					155				160	
Thr	Pro	Ala	Gln	Ile	Val	Ile	Ala	Gly	Glu	Val	Val	Ala	Val	Asp	Arg
			165						170					175	

B

Sub
a1

Ala Val Glu Leu Leu Gln Glu Ala Gly Ala Lys Arg Leu Ile Pro Leu
 180 185 190
 Lys Val Ser Gly Pro Phe His Thr Ala Leu Leu Glu Pro Ala Ser Gln
 195 200 205
 Lys Leu Ala Glu Thr Leu Ala Gln Val Ser Phe Ser Asp Phe Thr Cys
 210 215 220
 Pro Leu Val Gly Asn Thr Glu Ala Ala Val Met Gln Lys Glu Asp Ile
 225 230 235 240
 Ala Gln Leu Leu Thr Arg Gln Val Lys Glu Pro Val Arg Phe Tyr Glu
 245 250 255
 Ser Ile Gly Val Met Gln Glu Ala Gly Ile Ser Asn Phe Ile Glu Ile
 260 265 270
 Gly Pro Gly Lys Val Leu Ser Gly Phe Val Lys Lys Ile Asp Gln Thr
 275 280 285
 Ala His Leu Ala His Val Glu Asp Gln Ala Ser Leu Val Ala Leu Leu
 290 295 300
 Glu Lys
 305

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 450 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTGAAGAAAT CCTAAAAGAT TTGTATTACG GAGCCGCTAA GAAAATTCAA GAAGAAGCCT	60
CTCGCTGGGC AGGAGTTGTA AGAAATGACT AAAACAGCCT TTTTATTTGC TGGTCAAGGT	120
GCCCAGTATC TAGGGATGGG ACGGGATTTT TATGATCAGT ATCCGATTGT CAAAGAAACG	180
ATTEATCGAG CGAGTCAGGT GCTCGGTTAT GATTTACGTT ATCTCATCGA TACGGAAGAA	240
GACAAACTCA ATCAGACCCG CTATACGCAA CCAGCCATTC TAGCGACTTC GGTTGCTATC	300
TACCGTTTAT TGCAAGAAAA GGGCTATCAG CCTGATATGG TTGCTGGTTT GTCTCTTGGA	360

GAATACTCTG CCTTGGTGGC AAGCGGCGCC TTGGATTTTG AAGATGCGGT TGCCTTGCTA 420
 GCTAAGCGTG GAGCCTATAT GGAAGAAGCG 450

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 122 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Thr Lys Thr Ala Phe Leu Phe Ala Gly Gln Gly Ala Gln Tyr Leu
 1 5 10 15
 Gly Met Gly Arg Asp Phe Tyr Asp Gln Tyr Pro Ile Val Lys Glu Thr
 20 25 30
 Ile Asp Arg Ala Ser Gln Val Leu Gly Tyr Asp Leu Arg Tyr Leu Ile
 35 40 45
 Asp Thr Glu Glu Asp Lys Leu Asn Gln Thr Arg Tyr Thr Gln Pro Ala
 50 55 60
 Ile Leu Ala Thr Ser Val Ala Ile Tyr Arg Leu Leu Gln Glu Lys Gly
 65 70 75 80
 Tyr Gln Pro Asp Met Val Ala Gly Leu Ser Leu Gly Glu Tyr Ser Ala
 85 90 95
 Leu Val Ala Ser Gly Ala Leu Asp Phe Glu Asp Ala Val Ala Leu Val
 100 105 110
 Ala Lys Arg Gly Ala Tyr Met Glu Glu Ala
 115 120

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 547 base pairs
 (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TGCTGAACTT GGCAAGATGG TAGCAGTTCT CAATACGCCA GTAGAGGTCA TTGAAGAAGC 60
 CTGTCAAAAG CTTCTGGAAC TTGGAGTGGT TACTCCAGCC AACTATAACA CACCTGCACA 120
 AATCGTCATT GCTGGAGAAG TGTTGTCAGT TGATCGAGCG GTTGAAC TTTGCAAGAAGC 180
 AGGTGCCAAA CGCTTGATTC CTCTTAAGGT GTCAGGTCCC TTTCACACCG CTCTCCTTGA 240
 GCCTGCTAGC CAGAACTAG CTGAAACTCT AGCTCAGGTA AGTTTTTCAG ATTTTACTTG 300
 TCCCCTAGTC GGCAATACAG AAGCTGCTGT GATGCAAAAA GAGGACATTG CTCAGCTCTT 360
 GACGCGTCAG GTCAAGGAAC CCGTTCGTTT CTATGAAAGT ATTGGGGTCA TGCAAGAAGC 420
 AGGCATAAGC AACTTATTCG AGATTGGACC GGGGAAAGTC TTGTCAGGTT TTGTTAAAAA 480
 AATTGATCAA ACTGCTCACT TAGCTCATCT GGAAGATCAA GCGAGTTTAG TAGCACTTTT 540
 AGAAAAA 547

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 186 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Ala Glu Leu Gly Lys Met Val Ala Val Leu Asn Thr Pro Val Glu Val
 1 5 10 15
 Ile Glu Glu Ala Cys Gln Lys Leu Leu Glu Leu Gly Val Val Thr Pro
 20 25 30
 Ala Asn Tyr Asn Thr Pro Ala Gln Ile Val Ile Ala Gly Glu Val Val
 35 40 45

Ala Val Asp Arg Ala Val Glu Leu Leu Gln Glu Ala Gly Ala Lys Arg
 50 55 60
 Leu Ile Pro Leu Lys Val Ser Gly Pro Phe His Thr Ala Leu Leu Glu
 65 70 75 80
 Pro Ala Ser Gln Lys Leu Ala Glu Thr Leu Ala Gln Val Ser Phe Ser
 85 90 95
 Asp Phe Thr Cys Pro Leu Val Gly Asn Thr Glu Ala Ala Val Met Gln
 100 105 110
 Lys Glu Asp Ile Ala Gln Leu Leu Thr Arg Gln Val Lys Glu Pro Val
 115 120 125
 Arg Phe Tyr Glu Ser Ile Gly Val Met Gln Glu Ala Gly Ile Ser Asn
 130 135 140
 Leu Phe Glu Ile Gly Pro Gly Lys Val Leu Ser Gly Phe Val Lys Lys
 145 150 155 160
 Ile Asp Gln Thr Ala His Leu Ala His Val Glu Asp Gln Ala Ser Leu
 165 170 175
 Val Ala Leu Leu Glu Lys
 180